Environmental regulation and competitiveness

Florax, R.J.G.M.; Mulatu Gessesse, A.; Withagen, C.A.A.M.

2001

TI 2001-039/3
Tinbergen Institute Discussion Paper

Environmental Regulation and Competitiveness

Abay Mulatu[1,2]
Raymond J.G.M. Florax[1]
Cees A.A.M. Withagen[1,2,3]

[1] *Department of Spatial Economics, Master-point, Faculty of Economics and Business Administration, Vrije Universiteit Amsterdam,*
[2] *Tinbergen Institute,*
[3] *Department of Economics, and CentER, Tilburg University*

ENVIRONMENTAL REGULATION AND COMPETITIVENESS:

A Meta-Analysis of International Trade Studies

Abay Mulatu
Department of Spatial Economics
Master-point and Tinbergen Institute
Vrije Universiteit Amsterdam
De Boelelaan 1105
1081 HV Amsterdam
The Netherlands
Phone: + 31 20 4446049
Fax: + 31 20 4446004
E-mail: amulatu@econ.vu.nl
URL: http://www.econ.vu.nl/re/master-point

Raymond J.G.M. Florax
Department of Spatial Economics
Master-point
Vrije Universiteit Amsterdam
E-mail: rflorax@econ.vu.nl

Cees A.A.M. Withagen
Department of Spatial Economics
Master-point and Tinbergen Institute
Vrije Universiteit Amsterdam
and
Department of Economics and CentER
Tilburg University
E-mail: cwithagen@econ.vu.nl

Key words: environmental policy, strictness, competitiveness, meta-analysis
JEL code: F18, H30, L50, Q00

Abstract: The potential relationship between domestic environmental regulation and international competitiveness has evoked various speculations. The common neoclassical train of thought is that strict environmental regulation is detrimental to the competitiveness of industry, and that it induces phenomena such as ecological dumping, ecological capital flight, and regulatory 'chill' in environmental standards. A different view is that strict environmental regulation triggers industry's innovation potential, and subsequently increases its competitiveness. The impact of environmental regulation on competitiveness has been analyzed in terms of international capital movements, new firm formation, and international trade. This paper focuses on a statistically rigorous analysis of international trade studies, using a technique that is known as meta-analysis. The paper presents a statistically supported evaluation of the literature, in order to assess what the main conclusions regarding the relationship between environmental regulation and competitiveness are when it comes to studies on international trade flows. The synthesis of the literature is subsequently used to present guidelines for future primary research in this area.

1. INTRODUCTION

The persistent integration of the world economy has increased the apprehension for potentially negative effects that domestic environmental regulation may exert on a country's position in international trade. International trade and the environment are intertwined in various ways (Ulph 1997), among other things because the occurrence of international trade implies changing patterns of production and consumption that may have an impact on the level of pollution. Environmental pollution may also be the cause of spatial externalities, for instance, when consumption and production in one country imply non-negligible side effects on the level of pollution in other countries. Finally, the regulation of international trade through trade policies may be used to enforce international agreements on environmental issues.

The economic literature on international trade and the environment centers on the problems that arise from these relationships. It is sometimes argued that, in order to advance international trade and protect domestic firms at the same time, governments should not choose a stringent environmental policy, because a stringent policy may harm the competitiveness of domestic firms. A lenient environmental policy may however cause 'too much' pollution. Leniency can nevertheless be attractive as it prevents 'ecological capital flight', firms leaving the country and establishing production plants in places with a less stringent environmental policy (see Rauscher 1997 on the concept of ecological dumping). A contrasting view is purported in the hypothesis suggested by Porter (1991), who argues that a strict environmental policy may actually increase industries' competitiveness because it triggers innovation efforts of firms and reduces X-inefficiency.

These considerations are obviously interesting in the context of international trade theory, but they also have important policy implications. An appealing example of the latter is the introduction of a carbon tax, and the arguments this provoked in the political debate in the Netherlands. Several large energy consumers have, at least temporarily, been exempted from this tax because the legislator was persuaded by the argument that the Netherlands would otherwise become a less attractive location for industry. A close look at the economic literature shows, however, that only a fairly limited number of empirical studies addresses the issue of potentially negative effects of environmental policy on competitiveness measured in terms of trade flows.[1] These studies are mainly concerned with the United States.

The objective of the present paper is to review the empirical literature on environmental policy and trade and to assess its main findings. The relationship between domestic environmental policy — in particular its 'strictness' — and a country's international competitiveness is investigated in three major strands of the literature. The first set of studies deals with the impact of environmental regulation on international capital movements or foreign direct investment (see, e.g., Bouman 1998;

[1] Competitiveness has also been measured in terms of foreign direct investments and new firm formation (see below). The term 'competitiveness' is actually rather fuzzy, and has been interpreted and operationalized in different ways (for a general discussion, see Ekins and Speck 1999; Thomson 1998).

Hettige et al. 1992; Leonard 1998). The foreign direct investment literature generally reveals that empirical evidence on the ‘pollution haven’ hypothesis is fragile and inconclusive (Jeppesen et al. 2000). In a similar vein a second set of studies focuses on the impact of heterogeneous environmental regulation on the flow of capital in terms of domestic new firm formation. Although initially empirical evidence was thought to be rather weak, some recent studies have shown that environmental regulation affects the location behavior of pollution intensive manufacturing firms (Henderson 1996; Greenstone 1998; Becker and Henderson 2000; List and McHone 2000). In the third strand of literature, there is a rather limited set of studies that deals with the impact of environmental policy on international trade. Literature surveys of the latter (Jaffe et al. 1995; Jeppesen et al. 2000) point out that the available evidence for negative effects of environmental regulation on international trade is again rather mixed.

In what follows we will focus on the third strand of literature. We will use a traditional literature review as well as a statistically based approach, generally referred to as ‘meta-analysis’, to address the question whether the literature on environmental regulation and competitiveness is conclusive or not. The meta-analysis also serves as a useful starting-point for further primary research into the link between environmental policy and competitiveness, because insights derived from existing studies are of pivotal importance for the development of guidelines for future research.

The outline of the remainder of the paper is as follows. Section 2 introduces meta-analysis as an alternative approach to summarizing and synthesizing research results. Meta-analysis has been developed in the context of experimental sciences, and it constitutes a statistically rigorous approach to the assessment of research results. In Section 3 we review the literature, which constitutes the traditional approach to summarizing the state of the art in research. In this Section we also elaborate on the database employed in subsequent sections. In Section 4, various exploratory meta-analytical techniques are applied to the available set of studies on environmental policy and international trade flows. Section 5 presents a meta-regression analysis that systematically explains the variation in empirical results reported in the literature. The final section contains conclusions regarding the conclusiveness of the literature. It also elaborates on the implications of the meta-analysis for environmental policy-making, and presents some guidelines for future primary research into the relation between environmental policy and competitiveness.

2. STATISTICAL RIGOR THROUGH META-ANALYSIS

In a nutshell the empirical literature on the effect of stringency of environmental regulation on competitiveness is commonly characterized as one with mixed and rather vague evidence. This has led to divergent views on the issue, and numerous speculations as to the likely reasons for the differences and the lack of robustness of the results.

This conclusion is, however, based on a qualitative and narrative summary of what we think is the essence of the quantitative empirical results. This is the general practice in state of the art literature reviews compiled for journals and books. Although this practice is valuable in its own right, there are a number of problems associated with it. Obviously, the selection of the most relevant conclusions is to a certain extent arbitrary (Van den Bergh et al. 1997). Moreover, usually some sort of vote-counting procedure is implicit in a literature review. ‘Vote-counting’ refers to the practice of counting significantly positive, significantly negative and insignificant results. The statistical inference is that the category representing the majority of cases represents the true underlying relationship (Light and Smith 1971). Hedges and Olkin (1980) have shown that the vote-counting methodology is inadequate, because it tends to lead to the wrong conclusion more often when the number of studies increases. The basic argument is that the Type-II errors of the underlying studies do not cancel one another (see also Hedges and Olkin 1985). Notwithstanding this basic flaw of the vote-counting methodology we will show to what conclusion it leads when applied to the environmental regulation and international trade literature.

The rather crude comparison that is being made in vote-counting techniques (i.e., a categorical classification into (significantly) positive, zero, and negative effects) is also unsatisfactory because it is insufficient to determine whether the results of different studies agree (Hedges 1997). The difference in magnitude of the coefficients found in the literature should obviously be taken into account as well. Moreover, the results of an empirical study may provide a relatively good estimate of the sampling uncertainty of results, but non-sampling issues such as research design, model specification, and estimation techniques, are usually relatively constant within a study (Hedges 1997). Meta-analysis, in which non-sampling characteristics can be taken into account as moderator variables, constitutes a useful complementary technique to synthesize research results.

Meta-analysis has been developed in the context of the experimental sciences and refers to the statistical analysis of research results of studies performed previously. In economics meta-analysis is gaining ground, for instance, in industrial economics (Button and Weyman-Jones 1992), labor economics (Jarrell and Stanley 1990; Card and Krueger 1995; Ashenfelter et al. 1999), and transport economics (Button and Kerr 1996). Especially in environmental economics, stimulated by the work of Smith (1989), Smith and Kaoru (1990a,b), Smith and Osborne (1996) and Rosenberger et al. (1999), many meta-analyses appeared. The majority of the meta-analyses in economics are based on the so-called meta-regression technique. A meta-regression is usually based on least square estimation of a model in which a specific effect measure observed in a series of studies is taken as the dependent variable. The set of explanatory variables frequently includes specific underlying causes for the phenomenon under consideration, and moderator variables representing, for instance, differences among research designs, time-periods, and locations covered in the original studies (see, e.g., Stanley and Jarell 1989).

The advantage of meta-analysis over the more traditional literature review is obviously that it contributes to summarizing relationships and indicators, comparing the effect of the use of different methods, and tracing factors that are responsible for differing results across studies. However, there are a number of (practical) difficulties and limitations as well (see, e.g., Cooper and Hedges 1994). One of the most restrictive difficulties, especially in the context of domestic environmental regulation and international trade, is the incomparability of results. The results in this literature are in part exploratory and lack an explicitly defined effect size measure. The results merely convey the direction of the relationship under scrutiny, or — even in the case of a regression analysis — the coefficients are heterogeneous due to non-uniform specifications, definition of outcome measures, measurement of the variable of interest, etc.

Given such heterogeneity in the literature we resort to different meta-analytical techniques, depending on the information available in the primary studies. Vote-counting procedures, being the least demanding although admittedly rather crude, will be applied to all types of studies. Subsequently, in an attempt to circumvent — at least partially — the crudity of the vote-counting procedure, we will use an ordered probit model to analyze the sign and significance level of the environmental regulation's impact indicators, taking into account the influence of various moderator variables.

3. A SURVEY OF THE EMPIRICAL LITERATURE

The empirical literature on environmental policy and international trade can be divided into three groups depending on the approach being used: an exploratory approach, the Leontief approach, or an econometric approach. Each of these approaches will be discussed below, using the traditional state of the art review tool to summarize the main findings. Moreover, we set the stage for the next section by discussing in some detail how the data from each study have been used for the meta-analysis. This necessarily involves some technical details, most of which are treated in appendix A. In particular we will present so-called effect sizes below. These effect sizes can be defined as mean standardized differences between control and experimental groups. To standardize there are several options. One might use the pooled standard deviation, or the standard deviation of each separate group. It should be noted that under the null hypothesis of no effect of stringency on trade performance the effect sizes are Student-distributed. The appendix presents a more formal approach. Alternative definitions of effect size that are used are correlations and differences in proportions. It is fairly straightforward to extract the required probabilistic information from the econometric studies. This is more difficult for exploratory and Leontief type studies, that will therefore be given somewhat more attention below.

3.1 EXPLORATORY STUDIES

One approach to investigating the effect of regulation on international trade patterns is to explore whether a shift in trade patterns of pollution-intensive industries from developed to developing countries has occurred, the underlying assumption being that the shift is due to less restrictive environmental standards in developing countries. In this type of studies the impact of environmental policy is captured by a stringent and not-stringent dichotomy.

Low and Yeats (1992) consider 6 years (1965, 1975, 1985-1988) and make a distinction between industrial countries (EEC(10) and North America) and all other countries (Eastern Europe, Latin America, South–East Asia and West Asia). In the construction of our database the industrial countries are the experimental group (meaning subject to strict environmental regulation) and the developing countries are the control group.

The first type of data used from this study are those giving the share of environmentally dirty goods originating in different regions. To give an example, in 1975 is 40% of total trade in environmentally dirty goods originated in EEC(10). The effect size is calculated as follows. For the period 1975-1988 we calculated the average changes in the shares of the two groups, industrial and all other countries, as well as the pooled standard deviation using the changes in the individual regions within the groups. This yields an estimate of the effect size of –1.972. The effect size multiplied by the expression $\sqrt{n_e n_c/(n_e+n_c)}$ involving the sizes n_e and n_c of the experimental and control group, respectively, has a t-distribution. The t-value is -1.708, implying that the probability that a t-value is smaller than –1.708 or larger than 1.708 equals 0.163. This is the $p-$value used in the database. The interpretation of this figure is that it is safe to state that there is no significant decrease of the share of industrial countries in dirty exports. The same procedure used for data for the periods 1965-1988 and 1965-1975, in order to allow for the possibility that the results differ according to the period reviewed. No significant negative effects are found.

Another type of data refers to the same years and the same regions, but they give the regions' dirty exports as share of total exports from that region. For these data we apply the same procedure as outlined above. For these data the conclusion of no significant negative effect applies as well.

Finally, the third type of data refers to the concept of Revealed Comparative Advantage (RCA), defined as the share of a specific industry in a country's total exports as a fraction of the share of the industry in total world exports. Low and Yeats produce data on the number of countries with revealed comparative advantage indices exceeding unity. This is done for 40 industries over two time periods, 1966-1968 and 1986-1988. We take the pulp and wastepaper industry as an example. In the first period there were 5 industrialized countries with an RCA exceeding unity and 7 other countries with an RCA exceeding unity. Hence 42% of the countries with a share larger than unity was industrialized. For the second period this amounted to 44%. The effect size for this case is therefore 2%. The estimated standard deviation is calculated using the total number of countries having an RCA exceeding unity. The procedure outlined above yields a test statistic of 0.151, with a $p-$value of 12%. In this example the null hypothesis that the effect size is zero, is not rejected. Some of the empirical results demonstrate a tendency towards developing countries specializing in 'dirty' industries. For instance, there has been a disproportionately large increase in the number of developing countries that develop a revealed comparative advantage in 'dirty' industries. Over the period 1966-1988 there was 14 percent increase of industrial countries with revealed comparative advantage in dirty industries. For developing countries the percentage is approximately three times higher.

Sorsa (1994) provides data on the share of environmentally sensitive goods in total exports, for 7 industrial countries as well as for the total developing world, for 1970 and 1990. Since the control group consists of one element only we have used the standard deviations in the experimental group to compose the estimated effect size. The test statistic is –2.218, implying a $p-$value of 6.8%: hence it is likely that industrial countries export less dirty goods over time. Sorsa also provides data on revealed comparative advantage (not per industry). Again the same approach is taken, yielding a test statistic of –1.384 with $p-$value 21.6%. Thirdly, there are data on the correlation between the share in world trade of environmentally sensitive goods and the share of environmental expenditures in gross domestic product. These are calculated for private expenditures (three countries) and total expenditures (also three countries). The correlation coefficients can be used for testing because $t = r\sqrt{df} / \sqrt{(1-r^2)}$, where *r* denotes the correlation coefficient, has a *t*-distribution with *df* degrees of freedom.

Finally, we have used data on the number of industries with a revealed comparative advantage. Of all industries having a revealed comparative advantage 73.7% were in industrial countries in 1970. And it was 60.5% in 1990. This gives a test statistic of –1.233 and a corresponding $p-$value of 21.8%. Sorsa reaches the same conclusions as Low and Yeats, but she also shows that some countries maintain or even increase their comparative advantage. Moreover, the correlation analysis of environmental expenditures and the share of environmentally sensitive goods in world trade show no significant relationship.[2]

In sum, the empirical evidence of exploratory studies is at best mixed. It is obvious, however, that the exploratory studies suffer from various methodological deficiencies. Although they result in some interesting preliminary insights into the relationship between environmental regulation and international trade, they lack a theoretical basis, and use a poor specification of the differential effects of environmental regulation. Finally, an important drawback of this type of study is the inability to control for other factors that are potentially relevant to the observed changes in specialization patterns. For example, an increased demand for 'dirty' goods in developing countries can also be a cause for a production shift from developed towards the developing countries (see Jaffe et al. 1995).

[2] With the exception of Austria, for which Sorsa (1995) finds a significant effect, but this effect is positive.

3.2 *The Leontief approach*

A second approach towards assessing the regulation–competitiveness issue is in the spirit of Leontief's attempt to measure whether American exports are labor- or capital-intensive relative to imports. On the basis of Leontief's input-output model the pollution content of products can be assessed, taking into account the pollution related to direct inputs as well as to intermediate inputs from other sectors. Walter (1973) and Robison (1988) have used this approach. Also part of Kalt's (1988) work falls in this category. The theory behind the Leontief approach is based on Baumol and Oates (1975), who indicate that: "Undertaking pollution abatement will reduce the abating country's comparative advantage in producing high-abatement-cost goods and improve the comparative advantage in low-abatement-cost goods" (Robison 1988, p. 188). In the Leontief studies use is made of input-output matrices to calculate the overall abatement-cost contents of imports and exports.

Walter (1973) investigates the pollution content of US trade. Basically, the approach is as follows. A group of 83 goods and services is selected, and it is assumed that the share of environmental control costs in the final value of the goods is a proxy for the costs incurred to meet environmental criteria. Direct environmental production costs as well as indirect environmental costs[3] result in a measure labeled 'Overall Environmental Control Loading' (OECL). Subsequently, the OECL is multiplied by the value of US exports and imports to obtain the total environmental cost content of US trade. The ratio of abatement contents of imports to exports equals 0.81. The pollution contents of US exports and imports across all goods and services during 1960-1970 turns out to be 1.75% and 1.52% of total exports and imports, respectively[4]. Walter argues that the difference is negligible, and concludes that US environmental policy is generally trade neutral.

It is one of the aims of this section to extract probabilistic claims from each of the original studies. For the case at hand this is difficult. It contains no numerical comparison over time, so that nothing can be said about increasing stringency. There are however ratios of abatement contents for imports over exports in trade with Japan and Canada, amounting to 1.11 and 1.29 respectively. Treating the data as independent, which strictly speaking they are not, assuming that Japan and Canada have similar environmental policies we arrive at a positive impact of stricter environmental policy, but the effect is not significant.

Robison (1988) sets out to determine whether environmental control costs affected US comparative advantage, and what the impact of environmental cost is on US trade with Canada and the rest of the world. Input-output tables are used to determine the abatement cost content of US trade for the years

[3] The latter are calculated multiplying the abatement cost vector by the total requirement matrix.
[4] The import figure is obtained assuming that the pollution content of import commodities is the same as for domestically competing products.

1973, 1977 and 1982. Prices are endogenously determined through a full-fledged input-output model with 78 sectors. Abatement cost changes are assumed to be reflected in the value added of sectors, which subsequently translates into price changes via the input-output price equation. It is assumed that these price changes actually occur, at least in the long run. Abatement costs are defined in a way similar to Walter (1973), although Robison considers a 'modified' total requirement matrix in order to take account of abatement costs in capital goods. From the study we extract three types of data.

First Robison provides data for three years on the ratios of average abatement content of imports over exports, for total US trade and Canada-US trade. Over the three years the means are 1.236 and 1.115 with standard deviations of 0.133 and 0.029 respectively. The pooled standard deviation is 0.10, yielding an effect size of –1.249. The value of the test statistic is –1.529 with $p-$value of 20.1%. With the same data one can have a look at the differences in average annual change over time. This approach gives rise to a $p-$value of 41.2%.

Robison also considers the effect of a one percent price increase on trade for three years. We consider the average impact over the years, for total US trade as well as for US-Canada trade. We arrive at test statistics of –3.068 and -3.876 ($p-$values 9.2% and 6.1%, respectively), indicating that on average, over the years, the balance of trade deteriorates as a consequence of stricter environmental regulation.

Finally, Robison provides sectoral data on the impact of a one percent price increase on the balance of trade in the three years (in general as well as to and from Canada). We record the average and perform a test on the total effects over the sectors, per year. This procedure results in effect sizes that are significant and negative.

Part of Kalt's (1988) study employs an approach similar to Walter and Robison. For 1967 and 1977, Kalt provides data on the total abatement cost component of exports and of total abatement cost content of exports, for manufacturing industries only as well as for all industries. Moreover, he provides data on the factor dollars of abatement costs per dollar of exports and per dollar of imports for the same years and for both groups of industries. It is assumed in Kalt's set-up that the same abatement cost structure applied to both years. These data allow us to test in the usual way for changes in the abatement cost ratio over time in both types of industries. Kalt concludes that there is an indication that "environmental regulation was a source of shifting comparative advantage". However, our statistical analysis can not confirm this statement. Our test statistics for the import export ratio of abatement costs give values of –4.622 and –3.742 for all industries and manufacturing industries respectively, which are statistically not significant ($p-$values are 13.6% and 16.6%, respectively).

Much in the same way as for the exploratory approach, the studies based on the Leontief approach suffer from lack of conclusiveness. Although both Walter (1973) and Robison (1988) cast their conclusions in terms of environmental regulation, there are alternative — equally plausible — explanations for the observed small shifts in trade patterns. One frequently cited example is the process of industrialization in developing countries, which raises the relative importance of manufacturing in the economy (see, e.g., Jaffe et al. 1995).

3.3 ECONOMETRIC STUDIES

The econometric studies are based on either the Heckscher-Ohlin or on the gravity model. A brief description of the models is in order.

The Heckscher-Ohlin model rests on the following assumptions: factor immobility between countries, perfect factor mobility among industries, identical technologies in all countries, and different endowments of productive factors. It suggests that a country specializes in the production of commodities that require intensive use of resources that are relatively abundant (see, e.g., Helpman and Krugman 1985). Environmental policy can be easily incorporated in this kind of analysis: regulations deprive industries of the right to pollute, and can hence be considered a drain on endowments resulting in loss of comparative advantage.

The gravity model is frequently used in economics to model bilateral trade flows (see e.g., Helpman and Krugman 1985). Trade flows are specified as a function of the potential supply of the exporting country (measured by, e.g., Gross Domestic Product or population), the potential demand of the importing country (usually measured in a similar way), and some measure of friction to trade between the countries (oftentimes based on a distance measure). In order to investigate the impact of differing environmental regulations, an operational environmental variable is usually added to the set of variables that reflect the three factors mentioned above.

Kalt’s (1988) analysis is based on the Heckscher-Ohlin model. He specifies net exports of an industry as a function of physical capital, human capital, unskilled labor, research and development, and environmental control costs. He analyzes a cross-section of 78 industries for 1977, distinguishing three groups of industries: all industries, manufacturing and manufacturing excluding chemicals. For each cross-section two regressions are run, one with and one without correction for heteroscedasticity. A significant negative estimate of the environmental cost variable (costs of regulation based on a survey of 48 firms) is obtained only when the sample is confined to the 52 manufacturing industries. A similar result shows up for a specification where the change in net exports over the period 1967-1977 is the dependent variable, assuming that environmental costs were negligible in 1967. A third set of regressions involves net export performance in 1977 as the dependent variable and total direct and indirect factor inputs as independent variable. Pollution abatement expenditures have a significantly negative impact on export performance. Kalt concludes that environmental regulation had a significant negative effect on US manufacturing. As for our database, all coefficients are reported with their standard errors, and can therefore readily be incorporated into the database.

Tobey (1990) is concerned with multi-factor and multi-commodity extensions of the Heckscher-Ohlin model for the US. Net exports of each of 5 commodity groups, which are all qualified as dirty[5], are regressed on country characteristics, in particular endowments of land, capital, labor, natural resources (such as coal and oil production), and stringency of environmental regulation. The analysis includes trade flows to and from 21 countries, in 1975. Stringency of environmental regulation is represented by a qualitative index that ranges from 1 to 7 based on Walter and Ugelow (1979). The average of this index for developed countries is 6.1, and for developing countries the average equals 3.1. All estimates for the stringency variable turn out not to be statistically significant. Subsequently an omitted variable test is performed. It consists of two regressions; in the first regression strictness is not incorporated as an independent variable, in the second it is. If environmental policy does not play a role one would expect that half of the residuals is negative, and yields the same percentage for industrial and non-industrial countries. Therefore a test is performed on proportions. This is done for three groups: industrial/moderately developed, industrial/less developed and industrial/moderately plus less developed. The resulting differences in proportions are used in the meta-analysis.

Subsequently two extensions of the model, one allowing for non-homothetic preferences and one allowing for scale economies (larger countries having an advantage on the export market, reflected in national income as a fraction of world income as an explanatory variable), are considered. Regarding the first extension Tobey only mentions that environmental variable is not significant and that the omitted variable test does not support the hypothesis of a negative effect either. For the second extension no numerical data are reported either, but it is put forward that the stringency is not significant in the regression; the outcome of the omitted variable test is significant but points in the 'wrong' direction.

Tobey performs a second set of regressions, where the dependent variable is the change in exports in 1984 compared to 1970, and where the independent variable is the stringency index. This specification is motivated by the fact that differential stringency across countries in the 1960s may not have been strong enough to show up in a cross-section Heckscher-Ohlin model. Although resource endowments are thus basically assumed constant this specification does not reveal any significant effect either.

In a recent paper Van Beers and Van den Bergh (2000) perform a gravity analysis with the 1975 data employed by Tobey, for five dirty sectors, and a country sample similar to Tobey's. In their model

[5] Industries are labeled 'dirty' if show pollution abatement costs greater or equal than 1.85% of total costs, in 1977. The number of industries considered is 34, grouped into five commodity groups (i.e., mining, paper, chemicals, steel, and non-ferrous metals).

stringency does not have a significant effect in the chemicals and steel industry, whereas its effect is significantly negative for mining and non-ferrous metal, and significantly positive for the paper industry. The estimates are elasticities and are well-suited for the meta-analysis.

Diakosavvas (1994) also follows Tobey's (1990) framework, but his analysis centers on agriculture. Ten agricultural commodities for 23 countries (including five less developed) over the period 1984-1986 are considered in a Heckscher-Ohlin based cross-country model. Net exports of a country (in dollar terms) are regressed on endowments of labor and capital, the environment, government policies, and stringency of environmental regulation (i.e., the Walter and Ugelow measure). The regression results for five out of ten commodities subject to stricter environmental regulation suggest that environmental policy indeed causes net exports to fall. This also holds for another set of regressions where the dependent variable is total exports. The obvious difference with Tobey's results may be due to the restriction to the agricultural sector, the more recent time period, and/or different definitions of the explanatory variables.

Van Beers and Van den Bergh's (1997) analysis is essentially based on the Tobey (1990) approach as well, although there are three noteworthy differences. First, the gravity model, which considers bilateral instead of multilateral trade flows, is used. Bilateral exports are regressed on land area, GDP and population (as a proxy for potential supply of one country as well as potential demand in the other), and strictness of environmental policy in both countries. The distance between countries and dummy variables for membership of the European Union and the European Free Trade Association are added. Second, different measures of environmental regulatory stringency are considered. Third, three types of bilateral trade flows are used as dependent variables: total bilateral trade, 'dirty' bilateral trade, and 'dirty' footloose bilateral trade flows. Two variants of 'output oriented' environmental stringency measures are considered: a broadly defined measure using seven environmental indicators (e.g., protected land area, and paper recycling rate), and a narrowly defined measure based on a subset of these indicators considered to better reflect private environmental costs.[6] The sample consists of 14 OECD countries and 9 developing countries in 1975 and 21 OECD countries in 1992. In a series of ten regressions statistically significant results (elasticities with plausible signs) are found for the regressions of total bilateral exports or 'dirty' footloose exports on a set of variables including the narrow stringency measure. A positive relationship is found between strictness and export performance in 1975. For all estimates *t*-values are provided.

Differences vis-à-vis the Tobey (1990) results may be caused by various factors: for instance, commodity disaggregation, as not all industries are equally susceptible to regulation stringency (e.g.,

[6] 'Output oriented' measures reflect tangible outcomes of stringency of regulation. Although Van Beers and Van den Bergh (1997) make substantial efforts to improve the measurement of environmental stringency, Co et al. (1999) point out some caveats in their measures.

depending on whether they are footloose or not), the use of a better stringency measure, and the use of the gravity model. As the latter allows for bilateral instead of multilateral trade flows, stringency differentials cancel out.[7]

Xu (2000) replicates the Van Beers and Van den Bergh (1997) analysis, although with a number of modifications. First, a different measure of regulatory stringency is used, based on the work by Dasgupta et al. (1995). Second, the sample comprises 31 UNCED-report countries (which range from highly industrialized to extremely poor), randomly sampled from a total of 145 countries; this differs markedly from the Van Beers and Van den Bergh (1997) sample of 21 OECD countries for 1992. Finally, the effects of macroeconomic and cyclical disturbances are removed from the export flow variables, something that is ignored in Van Beers and Van den Bergh (1997). Regressions are performed for three dependent variables: bilateral exports, bilateral exports of environmentally sensitive goods and bilateral exports of non-resource-based (footloose) environmentally sensitive goods. Another distinction is between regressions having no import tariff variables included among the independent variables and another where they are included. Regressions are also performed using maximum likelihood to correct for possible heteroscedasticity. Altogether we obtain 12 estimated elasticities with t-values.

In part the results sharply contrast those obtained by Van Beers and Van den Bergh, as Xu finds statistically significant positive coefficients for the environmental policy variable, implying that a strict environmental policy is beneficial to export performance. He argues that this departure from the earlier results can be explained by the use of a different set of environmental policy measures, the removal of cyclical fluctuations from the export figures, and/or the inclusion of developing countries in the sample. The latter argument is, however, not fully convincing, as one would expect a negative impact of regulatory action given the likelihood of observing diverging regulatory standards in North-South trade.[8] Han and Braden (1996) use a Heckscher-Ohlin model for 19 manufacturing sectors in the US, for the period 1973-1990. Net export of a sector is expressed as a function of that sector's use of factors of production and pollution abatement cost, all of which show sectoral as well as temporal variation, captured by fixed or random effects terms[9]. In some regressions there are the expenditures on abatement multiplied by the time variable ($t = 1$ for 1973) in addition to abatement costs. This

[7] It should be noted that counterintuitive results, such as positive effects of stringency on exports, were obtained as well.

[8] In a another paper Xu (1999) considers five 'dirty' sectors (i.e., wood, paper and printing, chemicals, non-metal, and metal) in 30 countries including most of the OECD countries, in 1988, and he employs the same stringency measure as before. The paper reports an insignificant effect of environmental regulation on the share of dirty industries in total value added of the manufacturing industry. This study is, however, not considered in the sequel, as it does not address the trade effect of environmental regulation.

[9] Panel techniques also constitute a way to circumvent the problem of unmeasured heterogeneity inherent in OLS regressions. More arguments for the use of panel data in this context are given in Co et al. (1999).

variable is included to take account of possible variations in net exports due to time. So, denoting abatement expenditures by AB, the right-hand side of the regression includes $\beta_5 AB_{it} + \beta_6 AB_{it} t$.

In the fixed effects model the time invariant intercept captures the sector-specific effects. For four years (1975, 1980, 1985 and 1989) and all 19 sectors under consideration abatement expenditures elasticities of net exports are presented. This poses a problem for the construction of the database. The regression coefficients give the change in net exports in terms of dollar values. The t-values refer to these coefficients. It is therefore not straightforward to find standard errors corresponding to the elasticities. In the appendix we discuss this problem and present an approximation.

A second set of data is produced from a time-series regression per manufacturing industry. This yields sector-specific β_5's as well as β_6's. Given that the time variable is also involved, we can derive the impact of abatement costs, in terms of changes in exports in dollar values, for each sector and for each year. Again the problem is to find the standard error of effect sizes, which involves the unknown correlation coefficient. We refer to the appendix for a detailed treatment.

Finally, there are data on the panel regressions. Four panel regressions are executed. A distinction is made between one way and two way models, the difference being that in the two way model there is a time trend. A second distinction is between fixed effects and random error models. In the latter there is an individual time-independent error term in addition to the usual one. These models generate data for all years. Again it is necessary to construct standard errors since these are not readily available.

Of the series of regressions the study reports on, the results of the panel regression model show significant negative effects of environmental regulation on net exports.

The results of Han (1996) partly overlap with those reported in Han and Braden. We focus here on distinguishing features of Han's study. He considers the effect of differential environmental stringency on international competitiveness across countries, using an environmental regulation index based on the "the ratio of the emission reduction due to regulation and the emission in the absence of regulation". Using a Heckscher-Ohlin framework a panel of 34 countries at every five years interval from 1975 to 1990 is estimated for nine sectors. Of the five environmentally sensitive goods sectors considered, only mining is found to show a statistically negative relationship between stringency and competitiveness. Two other environmentally sensitive goods sectors (paper and pulp, and chemicals) on the other hand show significant positive estimates for the stringency variable. The effect of stringency on net exports of non-polluting sectors is insignificant. For all estimators *t*-values are provided.

Grossman and Krueger (1993) use cross-section data in a reduced form model to assess whether pollution abatement costs in the US explain the pattern of Mexican specialization and trade. In particular, patterns of US imports from Mexico, and US foreign direct investments in Mexico are scrutinized. Three different patterns are considered: one are the 1987 patterns of US imports from Mexico; another, the 1987 patterns of US imports from Mexico that have entered under the offshore assembly agreement; and a third one, the sectoral pattern of value added by 'maquiladora' plants.[10] The explanatory variables are human capital share, physical capital share, and tariff rates as well as a stringency variable. In addition, an injury rate is included in some of the regressions as a proxy for the major costs of US labor protection laws in American manufacturing. With regard to the first two patterns mentioned above, the stringency coefficient is expected to be positive. A positive relationship between the size of pollution abatement costs (as a fraction of value added) in the US manufacturing industry and the scale of sectoral activity in Mexico (imports of manufactured goods from Mexico) is reported[11]. It should be noted that Grossman and Krueger's regressions are not readily comparable to the others.

3.4 CONCLUSION

Table 1 summarizes the studies discussed in this section. It describes the studies according to the type of study and it gives the type of data we have been able to extract from each study. There are 13 studies. The study by Kalt is partly econometric and partly of the Leontief type. There is a large variety of estimates in the studies. There are univariate estimates, mainly from the Leontief studies. These estimates are means. Then there are also correlations and standardized mean differences, which are bivariate, and finally we have estimates from econometric studies, giving dollar value effects β s, and elasticities.

Insert table 1 about here.

We observe a large difference across studies in number of observations. We will pay attention to this below, when discussing the significance of the results from the studies. There are large differences in theoretical approach, which also make a comparison rather difficult. Finally we would like to mention that Leontief and exploratory studies do not control for other dimensions that might have an impact on the effect of environmental policy on competitiveness.

[10] 'Maquiladoras' are foreign-owned firms (usually somehow based in the US), with most of them located at the Mexican side of the US-Mexican border.

[11] A third regression concerns whether American firms invest in 'maquiladoras' to avoid environmental regulatory costs. As this pertains to direct foreign investment rather than trade flows, these estimates have been excluded from our analysis.

4. VOTE-COUNTING AND COMBINING P-LEVELS

4.1 VOTE-COUNTING

Simple vote-counting amounts to counting the number of studies (or results within studies) yielding positive, zero and negative effects. Such a procedure can be criticized on several grounds, including the fact that sample size is not taken into account, the fact that the method does not allow determining by how much the winner is winning and the fact that the statistical power is small (Bushman 1994). We nevertheless present the verdict of such a procedure as a starting point for other types of analyses. It should also be pointed out that vote-counting procedures are in principle meant for combining independent estimates. We have included multiple estimates from most studies, which makes it harder to justify independence. One has to take into account the significance of the values at a level common to all studies. We have used a relatively high level of 10% based on a two-sided test. In the tables below we make a distinction between results on effect sizes according to the categorization outlined above. There is a multitude of variables that can be included, and hence figures that can be produced. Therefore a selection has been made, and for tables and figures we present just three dimensions.

First of all we have made the distinction between exploratory, Leontief and econometric studies. Kalt's study is partly classified as econometric and partly as Leontief. Figure 1.A below gives the number of zero, negative and positive results per study. The striking features of the figure are that the Leontief studies produce relatively many negative effects. Moreover, the share of insignificant results is high in all study types. There are many negative significant estimates compared to positive significant estimates. This does not support the Porter hypothesis.

Insert Figure 1.A about here.

A second distinction made is according to the way the effect sizes are calculated. As in table 1 a distinction is made between β s, mean differences standardized by the control group, Glass'delta, standardized mean differences standardized by pooled standard deviation (Hedges' *g*), differences of proportions (Hedges' *g*-proportion), elasticities, differences from means and correlations. There are 691 observations in total, most of them β s (505) and elasticities (103). The results are summarized in Figure 1.B. Results in terms of differences in means produce relatively many negative significant results again. Also with (pooled) standardized mean differences we obtain a relatively large share of negative significant results.

Insert Figure 1.B about here.

Finally we have taken into account that stringency is measured in various ways. It can be measured by abatement costs, according to the Walter-Ugelow measure, the broad definition in Van Beers and Van den Bergh (VBVDB broad), the narrow definition of these authors (VBVDB narrow), the World Bank survey (World Bank), the measure developed by Han (Han) and qualitative measures (Others). Figure 1.C. summarizes the results.

Insert Figure 1.C about here.

Both measures employed by Van Beers and Van den Bergh yield relatively many negative and significant results (over 30%).

4.2 COMBINING SIGNIFICANCE LEVELS

There are many long established statistical methods for combining significance levels. They all deal with the question how probability values from independent studies can be combined. The hypothesis being tested can be expressed as follows:

$$H_0 : \vartheta_j = 0, \quad j = 1,2,...,k$$

where ϑ_j is the effect size in study j and k is the number of studies. If the null hypothesis is rejected then at least one of the population studies has a nonzero parameter. In the case at hand the null hypothesis will be that none of the studies supports a significant impact of stringency of environmental regulation on trade performance. Positive estimates are taken as evidence for the null hypothesis that $\vartheta_j = 0.$ (Cooper and Hedges 1994). So, the alternative hypothesis is

$$H_A : \vartheta_j \leq 0, \quad j = 1,2,...,k, \quad \vartheta_j < 0, \text{ for at least one } j.$$

Testing this alternative hypothesis requires the assumption that the population effects represented by $\vartheta_j -$values all are in the same direction (Cooper and Hedges, 1994, p. 220).

We consider four representative popular methods to combine significance levels.

- The minimum $p-$method, developed by Tippett (1931), rejects the null-hypothesis that in all studies the effect size is zero if:

$$\text{Min } (p_1, p_2,..., p_k) < \alpha := 1-(1-\alpha^*)^{1/k}$$

where α^* is the predetermined significance level for the combined significance test. We use α^*=0.05.

- As a second method we consider the sum of z 's method. It is based on the sum of the $z(p_i)$'s, where $z(p_i)$ is the $z-$value associated with p_i, divided by its standard deviation:

$$\sum_{i=1}^{k} z(p_i) \Big/ \sqrt{k}$$

- The third method is the sum of logs method.

$$-2\sum^{k} \log(p_j)$$

Under the null hypothesis this statistic has a χ^2 distribution with $2k$ degrees of freedom.

- Finally, we use the logit method. The expression

$$-\sum^{k} \log(p_j/(1-p_j))[k\prod^{2} (5k+2)/3(5k+40]^{-1/2}$$

is approximately $t-$distributed with $5k+4$ degrees of freedom.

Like in the simple vote counting discussed in the previous section we can make a large number of comparisons, according to the categories we wish to distinguish. In order to limit the number of tables we have maintained the categories used in the previous section.

It can be seen from the upper rows of Table 2 that for all types of studies all four methods almost always reject the null hypothesis that the relationship between environmental stringency and trade performance is not significantly different from zero. Only the minimum $p-$method does not reject the null hypothesis for exploratory studies. The middle part of Table 2 shows that the null hypothesis is rejected by all methods if the results are reported in terms of β s, elasticities, and difference from the mean. It seems that the minimum $p-$method rejects the null hypothesis least often. This also holds for the case where we consider the different stringency measures (lower section of Table 2). In this case we also find that for studies with abatement costs as a stringency measure all methods reject the null hypothesis.

Insert Table 2 about here

5. META-REGRESSION ANALYSIS

5.1 INTRODUCTION

A logical question that may arise from the results of the previous section is: if indeed the literature suggests that there is on average a statistically significant relationship for the issue under discussion, how can we explain the differential results (evidence) displayed by the primary studies? As mentioned in the review of Section 3, there have been speculations as to the likely factors accountable for the divergent results. The main dimensions of variation in the literature can be categorized as follows.

a. theoretical and methodological approach

To indicate the theoretical basis we use HOMOD as a variable, indicating the Heckscher-Ohlin model (the use of the gravity model is the omitted category). In this category we also have the three types of studies: exploratory (EXPLOR), Leontief (LEONTIEF) and econometric. The econometric studies can be subdivided into three categories: those that report on dollar values (ECTRBETA), those that provide elasticities (ECTRELAS) and other studies, i.e. the omitted variable test, as the reference category. This yields 4 variables.

b. operational focus

Stringency differentials can be represented by abatement costs or by a categorical stringency indicator (such as Walter-Ugelow, Han, Van Beers and Van den Bergh), denoted by STRINDEX, or by qualitative measures primarily used in the exploratory studies based on a qualitative assumption of stringency differentials being present (STRQUAL). We use abatement costs as the reference category. We also make a distinction between multi- or bilateral trade flows, which are incorporated by means of TFBILAT. While the literature displays cross-section, time-series, and panel data analyses, it has been argued that the first two could be inappropriate. Co et al. (2000), for example, suggest that, since the issue is analyzing "the difference between trade flows in country j at time t and trade flows in country j at time $t+\phi$, as a function of the difference in country j's environmental regulations", panel data analysis is the appropriate technique to investigate the issue. We have five types of data: time series of industries, cross-section of industries, cross-section of countries, panel of industries and panel of countries. These are grouped along two dimensions. We make distinguish between cross–section (CROSSEC), panel (PANEL) and time series, where the latter is the omitted category. Another dimension relates to some of the studies focusing on an analysis of industry-type data (INDDATA), whereas others focus on an analysis of countries, which is the omitted category.

c. spatial, temporal and sectoral dynamics

The variable YEAR indicates the year to which the estimate of the effect size refers. An important feature of the empirical literature under review is that the time period considered varies both across studies and within studies. Differences in degrees of environmental stringency across countries can be expected to diverge or converge over time. Therefore, estimates of different time periods may show systematic patterns. TSPANY indicates the length of the period to which the research applies.

Coverage of less developed countries is captured with the variable representing the ratio of the number of less developed countries to developed countries at the origin of the trade flow (LDCRATOR). The environmental regulation-competitiveness linkage can be considered as an essentially North-South issue. Differential stringency is expected between North and South, and hence to the extent that there is an effect of stringency on export performance, it should show up when one considers North-South trade.

A good way to take the impact of differences in pollution intensity and the degree to which an industry or sector is resource-based[12] into account is to include the variables POLLINT and NONRESB for pollution intensive, non-resource-based industries/sectors. These are the sectors where the most pronounced effect of a strict environmental policy can be expected. The reference category for these two variables are sectors that are not exclusively pollution intensive or resource-based. So, it includes the complement, which for POLLINT consists of studies where only non-pollution intensive studies are included as well as studies where both pollution-intensive and non-pollution intensive are included. A similar reasoning applies to NONRESB. In the regression we have also tried to study the interaction effect of POLLINT and NONRESB, but this interaction variable was never significantly different from zero.

d. measurement and estimation issues

This category deals with the definition of the dependent variable: the level of exports (YLEVEL) or the balance of payments (YBALANCE). It also includes the type of estimator (ESTOLS, with the more sophisticated techniques as omitted category, and ESTHET for those estimators that take into account heteroscedasticity).

For quick reference we insert table 3 that gives the set of variables used in the subsequent analysis.

Insert Table 3 about here

The first three categories contain (mainly) core variables, whereas the last category comprises (mainly) control variables.

5.3 MODEL AND ESTIMATION

The econometric model we use to estimate the general form of the meta-regression equation is an ordered probit model (see e.g., Greene 1997) The use of this model can be motivated as follows. We

[12] With regard to the definition of resource-basedness we employ the classification given in United Nations (1982). This is encompassing for industries, but not for agriculture. For the latter sector we rely on verbal information obtained from an agriculturist.

have collected a large group of statements about the effect of environmental stringency on trade performance. In principle the results vary from very significant negative effects to very significant positive effects. This is due to the different characteristics of the studies. In the database we do have effect sizes and significance levels but it is convenient to reduce the number of possible outcomes to three categories, labeled 0 for negative significant estimates, 1 for non-significant estimates, and 2 for positive significant estimates. The model reads as:

$$Y_i^* = \beta' X_i + \varepsilon_i, \ \ \varepsilon_i \sim N(0,1)$$

$$Y_i = 0 \ \text{ if } \ Y_i^* \leq 0$$

$$Y_i = 1 \ \text{ if } \ 0 < Y_i^* \leq \mu$$

$$Y_i = 2 \ \text{ if } \ Y_i^* > \mu$$

Here Y_i^* denotes the effect in study i. This is something we do not observe in the database, where we make a distinction only in the three categories described above; the observed counterpart to Y_i^* is Y_i. The variance of ε_i is assumed to be 1.0 since as long as Y_i^*, β and ε_i, are unobserved, no scaling of the underlying model can be deduced from the observed data. The μ denotes a threshold level, which is determined in the estimation procedure. The probit model is characterized by the assumption that

$$\text{Prob}(y = 0) = \Phi(-\beta' x)$$

$$\text{Prob}(y = 1) = \Phi(\mu - \beta' x) - \Phi(-\beta' x)$$

$$\text{Prob}(y = 2) = 1 - \Phi(\mu - \beta' x)$$

where Φ is the normal cumulative distribution. In case β_j is positive for some j, then an increase in the corresponding explanatory variable shifts the probability distribution to the right and therefore the probability of finding a zero y declines. In the model we use, there are only dummy variables, implying that we can not give the usual interpretation of marginal changes. However, positive coefficients still indicate that including the variable will decrease the probability of finding a zero y.

In principle a rather large number of regressions can be run and reported on. However, we restrict ourselves here to a discussion of two models: one with all variables included, called the full model, and one with a number of variables excluded (stripped version). The stripped version is based on exclusion of variables that are not significant. However, in any case those variables that are of theoretical interest are maintained throughout. These variables are ECTRBETA, ECTRELAS, POLLINT, NONRESB, YEAR and TFBILAT. Both the full and the stripped version are estimated with and without correcting for heteroscedasticity. In the correction for heteroscedasticity we use the number of observations in a study as the correction factor in a multiplicative form.The justification for this choice is that estimates from primary studies that are based on a larger number of observations have a smaller variance and should therefore be given more weight.

The results are displayed in Table 4.

Insert Table 4 about here.

The common-sense conclusion that the literature is not conclusive is not appropriate. Our findings indicate that some aspects are relatively clear-cut, whereas others are still rather vague and in need of further detailed primary investigation. The conclusions we draw from the ordered probit analysis are as follows.

There is no major difference between econometric studies measuring the effect of stringency by means of βs, elasticities or an omitted variable test. Econometric studies, however, find significantly fewer negative effects as compared to exploratory and Leontief-type studies: in all regressions the corresponding coefficients are significant and negative. This does not mean that the research methodology per se is responsible for this phenomenon. One could even argue that given our reservations with regard to the latter studies, they tend to make an overestimated contribution to the policy debate on stringency and trade.

There is a slight indication that including pollution intensive industries increases the probability of finding a statistically significant and negative effect. This result does not appear to be significant in case a correction for heteroscedasticity is applied.

The degree of mobility, being footloose or resource-based, does not seem to matter.

There is again a slight indication that in studies dealing with the more recent past there are significantly fewer estimates of negative effects of environmental policy to be found in the literature. This may indicate that the occasionally hypothesized convergence process is actually occurring.

Measurement issues, such as the length of the period under consideration and whether the effect on competitiveness is measured as a level or as a balance variable do not seem to matter.

The way in which stringency is measured is quite relevant for the outcome. When it is measured as a qualitative index, and even more so in case it is measured as a categorical index variable, the tendency to find more negative effects of environmental policy is present. Measurement of stringency by means of a qualitative binary indicator, or a categorical index variable based on various underlying indicators such as in, e.g., Van Beers and Van den Bergh, is crude relative to more precise measurement by means of a continuous abatement cost variable. Our finding implies that studies employing the former measures produce more significant negative results than the latter. The measurement issue, which is abundantly mentioned in the literature, therefore needs careful attention and further research should be done.

Subsequently there is a series of issues (mostly measurement and estimation issues) giving robust evidence that they contribute to finding significantly more negative trade effects of environmental policy. This is the case if sectoral rather than country data are used, if cross-section and panel data rather than time-series data are used and if trade flows are measured bilaterally instead of multilaterally. It also holds when the evidence is obtained by OLS. A correction for heteroscedasticity does not really matter.

In line with what economic theory suggests the presence of less developed countries in the sample (at the origin of the trade flow) leads to significantly less negative effects to be detected.

Finally, the Heckscher-Ohlin framework, as opposed to the gravity model, induces the occurrence of negative effects of environmental policy.

In sum, the literature is not inconclusive in many respects, although some qualifications need to be made here, because for instance we derive these conclusions heavily relying on Han and Braden (1996) estimates. However, the pivotal issues of the measurement of stringency of environmental policy and the pollution intensity and resource basedness are yet unsettled as the results do not appear to be very robust with respect to these variables.

6. CONCLUSION

This paper is essentially a critical review of the empirical literature on environmental regulation and international trade flows, a literature that displays contradictory evidence.

It starts with the customary practice of outlining the salient features of each study and summarizing the results. The conclusion from this review, like similar review exercises, is that the empirical literature does not strongly support the hypothesis that the effect of environmental regulation on competitiveness is negative. The qualitative review identifies several controversial issues in the literature. Most prominent are the stringency measure, type of data, methods of investigation and sectoral disaggregation.

Four types of meta-analytic techniques of combing significance levels are applied to the studies under discussion. The results suggest that there is almost always at least one study/estimate in the available literature that displays a statistically significant negative relationship between environmental stringency and international trade flows. Stated otherwise, the hypothesis of no effect is rejected in almost all tests.

Furthermore, in an attempt to explain the possible factors for the divergent results, meta-regression analysis is used. The studies/estimates differ with respect to so many dimensions (data type, measure of effect size, etc.) that only in a multivariate context one gets a good idea of what the literature "tells" you. The results of the ordered probit analysis indicate that for the likelihood of a study to find a negative relation between environmental regulation and international trade flows, pollution intensity of the sectors involved is barely significant. The property that a sector is resource based has the "right" sign (mostly) but is not very significant either. Inclusion of abatement costs as a stringency measure reduces the probability of finding a significant and negative effect. Inclusion of developing countries in the sample increases the probability of finding a negative relationship.

What are the insights one can gain for future primary research in this area? The environmental regulation-competitiveness linkage should be investigated with data on the industrial level rather than on the country level. This is to be preferred from a theoretical point of view, because environmental policy is usually industry related rather than general. But it also follows from our findings that the hypothesis of no effect of stringency of environmental policy on trade performance should be tested at a low level of aggregation: Rejection of the hypothesis at that level seems to warrants rejection at the higher level as well. In addition, the fact that environmental stringency differentials are likely to prevail more between North and South, inclusion of developing countries in the sample is desirable. Also, since most of the studies under scrutiny did not involve the European countries, it seems worthwhile to have a much closer look at this region, in relation to developing countries. Similarly, the diverse measures of stringency are one of the factors for the disparate evidence in the empirical literature. Therefore, any new empirical investigation of the relationship under discussion should come to grips with this problem before embarking on estimation by picking up one or the other controversial measures.

REFERENCES

Ashenfelter, O., Harmon, C. and Oosterbeek, H. (1999): "A review of estimates of the schooling/earnings relationship, with tests for publication bias", *Labour Economics 6*, pp. 453-470.

Baumol, W. and Oates, W. (1975), *The Theory of Environmental Policy*, Prentice Hall, New Jersey.

Becker, R. and Henderson, V. (2000): “Effects of Air Quality Regulations on Polluting Industries”, *Journal of Political Economy* 108, pp. 379-421.

Beers, C. van, and Bergh, J. van den (1997): “An empirical multi-country analysis of the impact of environmental regulations on foreign trade flows”, *Kyklos 50*, pp. 29-46.

Beers, C. van, and Bergh, J. van den (1998): “The impact of environmental policy on foreign trade: an empirical multi-country analysis based on bilateral trade flows”, *mimeo*, Free University Amsterdam.

Beers, C. van, and Bergh, J. van den (2000): “The impact of environmental policy on foreign trade: tobey revisited with a bilateral trade flow”, *mimeo,* Department of Economics, Delft University of Technology, Delft.

Bergh, J. van den, Button, K., Nijkamp, P. and Pepping, G. (1997), *Meta-Analysis in Environmental Economics*, Kluwer Academic Publishers, Dordrecht.

Bouman M. (1998), *Environmental Costs and Capital Flight*, PhD thesis, Tinbergen Institute Research Series, University of Amsterdam, Amsterdam.

Bushman, B. (1994): “Vote counting procedures in meta-analysis”, in H. Cooper and L. Hedges (eds.), *The Handbook of Research Synthesis*, Russell Sage Foundation, New York, pp. 193-214.

Button, K. and Kerr, J. (1996): “Effectiveness of traffic restraint policies: a simple meta regression analysis”, *International Journal of Transport Economics 23*, pp. 213-225.

Button, K. and Weyman-Jones, T. (1992): “Ownership structure, institutional organization and measured X-inefficiency”, *American Economic Review 82, Papers and Proceedings*, pp. 439-445

Card, D. and Krueger, A. (1995): “Time-series minimum-wage studies: a meta-analysis”, *American Economic Review 85*, pp.238-243.

Co, C., Folmer, H., Jeppesen, T. and List, J. (1999): “Impact of environmental regulation on foreign trade flows: methodological problems and possible solutions”, *working paper*, CentER Tilburg University.

Cooper, H. and Hedges, L. (1994), *Handbook of Research Synthesis*, Russell Sage Foundation, New York.

Dasgupta, S., Mody, R., Roy, S. and Wheeler, D. (1995): “Environmental regulation and Development: a Cross-Country Empirical Analysis”, *World Bank Policy Research Working Paper No 1448.*

Diakosavvas, D. (1994): “The impact of environmental policies on agricultural trade”, *Journal of International Development 6,* pp. 207-218.

Ekins, P. and Speck, S. (1999): “Competitiveness and exemptions from environmental taxes in Europe”, *Environmental and Resource Economics 13*, pp. 369-396.

Greene, W. (1997), *Econometric Analysis*, McMillan, New York

Greenstone, M. (1998): “The Impacts of Environmental Regulations on Industrial Activity: Evidence from the 1970 and 1977 Clean Air Act Amendments and the Census of Manufactures,” *Working Paper #408*, Princeton University.

Grossman, G. and Krueger, A. (1993): “Environmental impacts of a North American free trade agreement,” in P. Garber (ed.), *The U.S.-Mexican Free Trade Agreement*, MIT-Press, Cambridge MA, pp.13-56.

Han (1996), *Environmental regulation and International Competitiverness: Cross-Industry and Cross-Country Analyses*, PhD-thesis, University of Illinois.

Han, K. and Braden, J. (1996): “Environment and trade: new evidence from U.S. manufacturing”, *mimeo*, University of Illinois.

Hedges, L. (1997): “The promise of replication in labour economics”, *Labour Economics 4*, pp. 111-114

Hedges, L. and Olkin, I. (1980): “Vote counting methods in research synthesis”, *Psychological Bulletin 88*, pp.359-369.

Hedges, L. and Olkin, I. (1985), *Statistical Methods for Meta-analysis*, Academic Press, NewYork.

Helpman, E. and Krugman, P. (1985), *Market Structure and Foreign Trade*, MIT Press, Cambridge MA.

Henderson, J.V. (1996) “Effects of Air Quality Regulation”, *American Economic Review 86*, pp. 789-813.

Hettige, H., Lucas, R. and Wheeler, D. (1992): “The toxic intensity of industrial production: global patterns, trends, and trade policy”, *American Economic Review 82, Papers and Proceeding,* pp. 478-481.

Jaffe, A., Peterson, S., Portney, P. and Stavins, R. (1995): “Environmental regulation and the competitiveness of US manufacturing: what does the evidence tell us?”, *Journal Economic Literature 33*, pp. 132-163.

Jarrell, S. and Stanley, T. (1990): “A meta-analysis of the union non-union wage gap”, *Industrial and Labor Relations Review*, 54-67.

Jeppesen, T., Folmer, H. and Komen, M. (2000): “Environmental policy, international trade and capital movement: a synopsis of the macroeconomic literature”, in Th. Sterner (ed.), *The Market and the Environment: The Effectiveness of Market-based Policy Instruments for Environmental Reform*, Edward Elgar, Cheltenham, pp. 83-113.

Kalt, J. (1988): “The impact of domestic environmental regulatory policies on U.S. international competitiveness”, in M. Spence and H. Hazard (eds.), *International Competitiveness*, Ballinger, Harper and Row, pp. 221-262.

Leonard, H. (1988), *Pollution and the Struggle for the World Product*, Cambridge University Press, Cambridge.

Light, R. and Smith, P. (1971): “Accumulating evidence: Procedures for resolving contradictions among different research studies”, *Harvard Educational Review 41*, pp. 429-471.

List, J. and McHone, W. (2000): “*Effects of Environmental Regulations on Manufacturing Plant Births: Evidence from a Propensity Score Matching Estimator*”, Paper presented at the European Science Foundation (EURESCO) meeting on the International Dimension of Environmental Policy, Kerkrade, The Netherlands.

Low, P. and A. Yeats (1992): “Do 'dirty' industries migrate?”, in P. Low (ed.), *International Trade and the Environment*, Washington DC, World Bank.

Phillips, J. and Goss, E. (1995): “The effects of state on local taxes on economic development: a meta-analysis”, *Southern Economic Journal 62*, pp. 320-333.

Porter, M. (1991): “America’s green strategy”, *Scientific American*, p.168.

Rauscher, M. (1997), *International Trade, Factor Movements, and the Environment*, Oxford University Press, Oxford.

Robison, H. D. (1988): “Industrial pollution abatement: the impact on balance of trade”, *Canadian Journal of Economics 1*, pp.187-199.

Rosenberger, R. and Loomis, J. and Shrestha, R. (1999): “Thirty years of outdoor recreation use value studies: Testing for trends and study factor effects using meta-analysis” *mimeo*, Department of Agricultural and Resource Economics, Colorado State University, Forth Collins, Colorado.

Smith, V. (1989): “Can we measure the economic value of environmental amenities?”, *Southern Economic Journal 56*, pp. 865-878.

Smith, V. and Kaoru, Y. (1990a): “Signals or noise: explaining the variation in recreation benefit estimates”, *American Journal of Agricultural Economics 72*, pp. 419-433.

Smith, V. and Kaoru, Y. (1990b): “What have we learned since Hotelling’s letter? A meta-analysis”, *Economics Letters 32*, pp. 267-272.

Smith, V. and Osborne, L. (1996): “Do contingent valuation estimates pass a “Scope” test? A meta-analysis”, *Journal of Environmental Economics and Management 31*, pp. 287-301.

Sorsa, P. (1994), *Environmental Protection, North-South Trade, and the Uruguay Round*, International Monetary Fund, Washington DC.

Stanley, H. and S. Jarrell (1989): “Meta-regression analysis: a quantitative method of literature surveys”, *Journal of Economic Surveys 3*, pp.161-170.

Thompson, G. (1998): “International competitiveness and globalization: framework for analysis, connections and critiques”, in T. Barker and J. Köhler (eds.), *International Competitiveness and Environmental Policies*, Edward Elgar, Cheltenham, pp.13-31.

Tippett, L. (1931), *The Methods of Statistics*, Williams&Norgate, London.

Tobey, J. (1993): “The impact of domestic environmental policy on international trade”, in H. Giersch (ed.). *Economic Progress and Environmental Concerns*, Springer Verlag, Berlin.

Ulph, A. (1997): “International trade and the environment: a survey of recent economic analysis”, in H. Folmer and T. Tietenberg (eds.), *The International Yearbook of Environmental and Resource Economics 1997/1998*, Edward Elgar, Aldershot, pp. 205-242.

United Nations (1982), *Changing Patterns of Trade in World Industries,* United Nations, New York

Walter, I. (1973): “The pollution content of American trade”, *Western Economic Journal 1,* pp. 61-70.

Walter, I. and Ugelow, J. (1979): “Environmental policies in developing countries”, *Ambio 8,* pp. 102-109.

Xu, X. (1999): “Do stringent environmental regulations reduce the international competitiveness of environmentally sensitive goods? A global perspective”, *World Development 27*, pp. 1215-1226

Xu, X. (2000): “International trade and environmental regulations: time series evidence and cross section test”, *Environmental and Resource Economics* 17, pp. 233-257

Appendix

The objective of this appendix is to introduce the so-called effect size used in the meta-analysis. The effect size generally refers to the effect of stringent environmental policy on export performance. Meta-analysis offers several methods to obtain the effect size. Two of these methods are amply used in the paper and will therefore be described in some detail in this section. We also elaborate on the way the data from the study of Han and Braden are used.

EFFECT SIZE

Generally a distinction is made between a control group, with index c, consisting of n_c elements, and an experimental group, with index e, consisting of n_e elements. Two alternative effect size measures are the Glass' Δ and Hedges's g defined as

$$\Delta = \frac{M_e - M_c}{S_c} \text{ and } g = \frac{M_e - M_c}{S_p}$$

respectively, where M_c and M_e are the means in the control group and the experimental group respectively and S_c and S_p are the standard deviations in the control group and the pooled standard deviations respectively. The latter is defined by:

$$S_p = \sqrt{\frac{((n_c - 1)S_c)^2 + ((n_e - 1)S_e)^2}{n_c + n_e - 1}}$$

where S_e is the standard deviation in the experimental group. In the meta-analytical literature it is argued that the pooled standard deviation is a better estimate of the population standard deviation, but choosing the standard deviation of the control group is deemed a "very reasonable alternative" ((Rosenthal 1994, p. 232)). Then we can use the fact that

$$\text{A1. } \Delta\sqrt{\frac{n_c n_e}{n_c + n_e}} \text{ and A2 } g\sqrt{\frac{n_c n_e}{n_c + n_e}}$$

have a t- distribution with $n_c + n_e - 2$ degrees of freedom, under the commonly made assumption of a normal distribution of the original observations.

When the data are proportions the effect size is defined by $p_e - p_c$, with a standard error defined by

$$SE^2 = \frac{p_e(1-p_e)}{n_e} + \frac{p_c(1-p_c)}{n_c}$$

We can then make use of the normal distribution to derive confidence intervals and $p-$values.

HAN AND BRADEN (1996)

The study by Han and Braden provides the majority of the data for our analysis. First, for 19 product groups and for four years (1975, 1980, 1985 and 1989) they present elasticities giving the percentage change in net exports following a one percent increase in abatement expenditures *AB*. These elasticities are the result of a regression equation of the type:

$$X_{it} = ... + \beta_5 AB_{it} + \beta_6 AB_{it} t + ...$$

where *t* refers to time. The elasticity is:

$$\varepsilon_{it} = \frac{\partial X_{it}}{\partial AB_{it}} \frac{AB_{it}}{X_{it}}$$

We are interested in finding the standard error of the estimate of the elasticity. To that end an assumption has to be made, because the data do not allow for a straightforward calculation. We have chosen to assume that in every year the ratio of abatement expenditures over net exports is the same, for all sectors. It is denoted by c. That implies that the standard error can be written as

$$c(Var(\beta_5) + t^2 Var(\beta_6) + 2t.rSE(\beta_5)SE(\beta_6))^{1/2}$$

In this expression r is the correlation coefficient and the β's refer to the estimates. The time variable assumes the values 3, 8, 13 and 17. (1975 is 3 because the initial data are from 1973). The variances of the individual estimates are given in Han and Braden. Next we had to chose values for c and r. To be on the safe side we took $c = 0.1$ and $r = 0.8$. This procedure enables us to perform the usual tests on effect sizes.

Han and Braden also present a cross-section regression over the industries for each year between 1973 and 1990. This yields another set of data on the effect of abatement costs on net exports. To each estimate there corresponds a $t-$value. Finally, there is also a time series analysis by manufacturing sectors. In this analysis for each sector there is an explanatory variable of the form AB_t and one of the form $AB_t t$. This leaves us with 18 estimates per sector. For both explanatory variables the $t-$values are given separately. In the two cases, panel and time series, we have applied the same procedure as outlined above to obtain the standard deviation of the impact of abatement costs on trade performance.

Table 1. A categorization of the stringency-competitiveness trade flow literature according to type of approach, type of estimate, and number of available estimates

Study	*Type of approach*[a]			*Type and number of estimates*[b]							
	Exploratory	*Leontief*	*Econometric*	*Univariate*	*Bivariate*				*Multivariate*		Total
				μ	r	g	g_p	Δ	β	η	
Walter (1973)		×		1							1
Kalt (1988)		×	×	4	4				18		26
Robison (1988)		×		8		2					10
Tobey (1990)[c]			×				9		20		29
Low and Yeats (1992)	×					6	40				46
Grossman and Krueger (1993)			×						2		2
Diakosavvas (1994)			×						20		20
Sorsa (1994)	×				6		1	2[d]			9
Van Beers and Van den Bergh (1997)[e]			×							10	10
Han (1996)			×						9		9
Han and Braden (1996)			×						436	76	512
Xu (2000)			×							12	12
Van Beers and Van den Bergh (2000)			×							5	5
Total	2	3	9	13	10	8	50	2	505	103	691

[a] The categories refer to exploratory, Leontief, and econometric approaches, as in Section 3. One study combines different approaches.
[b] The symbols used are: μ is a sample mean, r a correlation coefficient, g Hedges' difference in means, g_p Hedges' difference in proportions, Δ Glass' difference in means, β a regression coefficient not representing an elasticity, and η an elasticity (see Appendix A for details).
[c] Tobey (1990) presents 13 estimates in a quantitative fashion. The results of two alternative specifications are mentioned in the text, but no numerical details are given. Sufficient information for binary and categorical coding is, however, available.
[d] The standard deviation is based on the experimental group because too few observations are available for the control group (see Appendix A for details).
[e] Not all results are published in the *Kyklos* version of the paper; estimates are taken from an extended version of the article (Van Beers and Van den Bergh 1996).

Table 2. Results of combining P-values according to various categorizations

Typology of study	Method of combining P-values							
	Minimum P		Sum of z		Sum of logs		Logit	
	Minimum P	Critical level	Value of test-statistic	Probability	Value of test-statistic	Probability	Value of test-statistic	Probability
According to approach								
Econometric	0.000001	0.000085	-12.818	0.000	2325.800	0.000	15.543	0.000
Leontief	0.000003	0.002696	-8.043	0.000	178.964	0.000	12.669	0.000
Exploratory	0.006086	0.000932	-2.474	0.007	139.631	0.030	2.384	0.018
All studies	0.000001	0.000076	-14.151	0.000	2644.395	0.000	17.110	0.000
According to effect size								
Beta	0.000022	0.000104	-9.208	0.000	1434.926	0.000	9.947	0.000
Elasticity	0.000001	0.000498	-10.827	0.000	885.956	0.000	16.288	0.000
Correlation	0.071608	0.005116	0.805	0.790	16.141	0.708	-1.553	0.126
Mean	0.000003	0.003938	-9.098	0.000	165.386	0.000	15.526	0.000
Hedges'g	0.081347	0.006391	-2.019	0.022	25.077	0.068	2.601	0.013
Hedges' g-proportion	0.006086	0.001165	-2.091	0.018	105.707	0.096	2.213	0.028
Delta	0.034233	0.025321	-2.164	0.015	11.203	0.024	4.339	0.001
According to type of stringency								
Walter-Ugelow	0.004937	0.001314	-2.866	0.002	159.161	0.000	2.589	0.010
Abatement cost	0.000001	0.000092	-15.352	0.000	2294.155	0.000	19.324	0.000
Han	0.008941	0.005683	1.221	0.889	18.018	0.454	-1.696	0.096
World Bank	0.367030	0.004265	10.524	1.000	2.254	1.000	-18.236	0.000
Others	0.006086	0.000967	-2.881	0.002	136.430	0.025	2.985	0.003
VBVDB narrow	0.002093	0.016952	-3.124	0.001	24.497	0.000	6.574	0.000
VBVDB broad	0.031186	0.016952	-0.857	0.196	9.880	0.130	1.591	0.128

Table 3. List of variables

Variable	Description
EXPLOR	Exploratory study
LEONTIEF	Leontief type study
ECTRBETA	Econometric study that reports dollar values
ECTRELAS	Econometric study that reports elasticities
POLLINT	Estimate for pollution intensive sector
NONRESB	Estimate for non-resource based sector
TSPANY	Time-span of dependent variable
YEAR	The year to which the estimate reefers
YLEVEL	Dependent variable measured as a level
YBALANCE	Dependent variable measured as a balance
STRABAT	Stringency measured by means of abatement cost
STRQUAL	Stringency measured in a qualitative way
STRINDEX	Stringency measured by means of indices
INDDATA	Industry data
TIMESER	Time series data
CROSSSEC	Cross-section data
PANEL	Panel data
ESTOLS	OLS estimator
ESTHET	Estimator corrects for heteroscedasticity
HOMOD	Heckscher-Ohlin model
LDCRATOR	Ratio of number of developing countries to that of developed countries considered at the origin of the trade flow
TFBILAT	Bilateral trade flows

Table 4. Results for an ordered probit specification of four different versions

Variable [a]	FULL MODEL		RESTRICTED MODEL		FULL MODEL WITH HETEROSCEDASTICITY CORRECTION		RESTRICTED MODEL WITH HETEROSCEDASTICITY CORRECTION	
	Parameter	t-ratio	Parameter	t-ratio	Parameter	t-ratio	Parameter	t-ratio
Constant	-93.3521**	-2.4944	-83.0836**	-2.2193	-144.5900	-1.3025	-136.7290	-1.2751
EXPLOR	-2.8184**	-2.2931	-2.3402**	-2.3849	-4.0632*	-1.8693	-3.5869**	-2.1038
LEONTIEF	-4.0493**	-2.3801	-3.8307**	-2.9387	-5.6452**	-2.2499	-5.2541**	-2.6499
ECTRBETA	1.1876	0.6083	0.9051	0.5920	1.3686	0.5964	1.2818	0.6515
ECTRELAS	1.0634	0.5422	0.7652	0.5030	1.1414	0.4922	1.0433	0.5438
POLLINT	-0.4930*	-1.6912	-0.4603	-1.6022	-0.6546	-1.1147	-0.6283	-1.1302
NONRESB	0.0015**	0.0070	-0.0272	-0.1312	-0.0849	-0.2256	-0.1074	-0.2953
TSPANY	-0.2399	-0.3174			-0.1836	-0.2373		
YEAR	52.5105**	2.8018	47.0308**	2.5005	80.1253	1.4097	75.8014	1.3828
YLEVEL	-0.5515	-0.8245			-0.5729	-0.7116		
YBALANCE	0.0694	0.1705			0.0673	0.1341		
STRQUAL	-2.6706**	-2.6138	-2.9242**	-3.9038	-4.3770**	-2.6975	-4.5953**	-3.2076
STRINDEX	-5.2495**	-4.4644	-5.3717**	-5.5432	-7.5153**	-3.7422	-7.7041**	-4.1216
INDDATA	-4.7557**	-4.1439	-4.7392**	-4.9177	-6.5054**	-3.9287	-6.4530**	-4.4047
CROSSSEC	-1.9889*	-1.8757	-1.6263**	-4.3663	-2.2795**	-2.0726	-2.0322**	-4.8563
PANEL	-4.2817**	-6.8627	-4.5060**	-8.4377	-5.4247**	-4.6298	-5.4564**	-4.9443
ESTOLS	-2.3848**	-4.7578	-2.6427**	-7.6198	-3.1607**	-3.0639	-3.2196**	-3.5475
ESTHET	-0.2293	-0.6868			-0.0087	-0.0192		
HOMOD	-1.9558**	-2.3245	-1.4208**	-1.9785	-2.7653**	-2.1653	-2.2254**	-2.0557
LDCRATOR	1.8289**	9.4745	1.9088**	10.1022	2.8211**	3.9452	2.9153**	4.1201
TFBILAT	-1.2413*	-1.7184	-1.2528*	-1.8359	-1.1237	-1.3556	-1.0457	-1.4279
Mu(1)	4.3028	16.3746	4.2633	16.9923	5.5797	5.5141	5.5567**	6.0855
NOBS					0.2615	2.6047	0.2630	2.7520
Log-L	-235.321		-236.599		-220.146		-220.773	
Log-L(0)	-386.529		-386.529		-386.529		-386.529	

[a]The dependent variable is coded with 0 for negative effect sizes, 1 for effect sizes that are not significantly different from zero, and 1 for positive effect sizes
* = significant at 10 % level of significance.
** = significant at 5% level of significance.

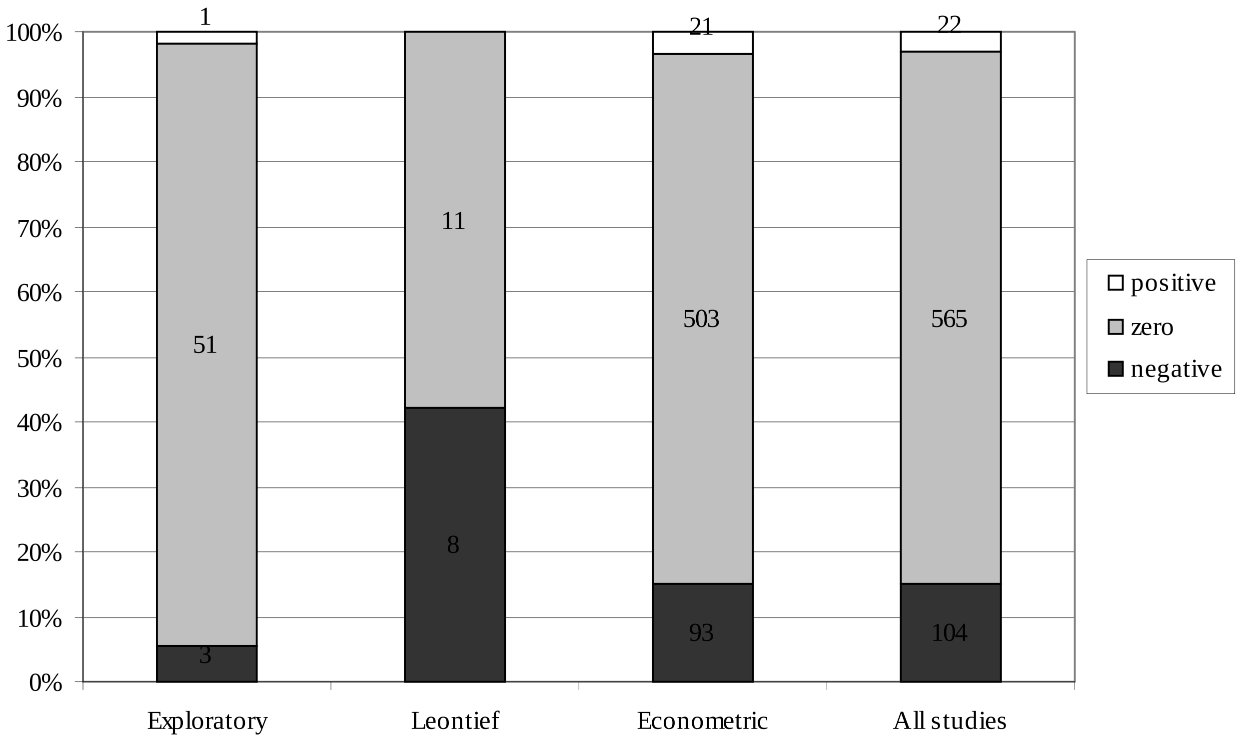

FIGURE 1. A. Percentage and number of negative, zero, and positive effect sizes according to study approach

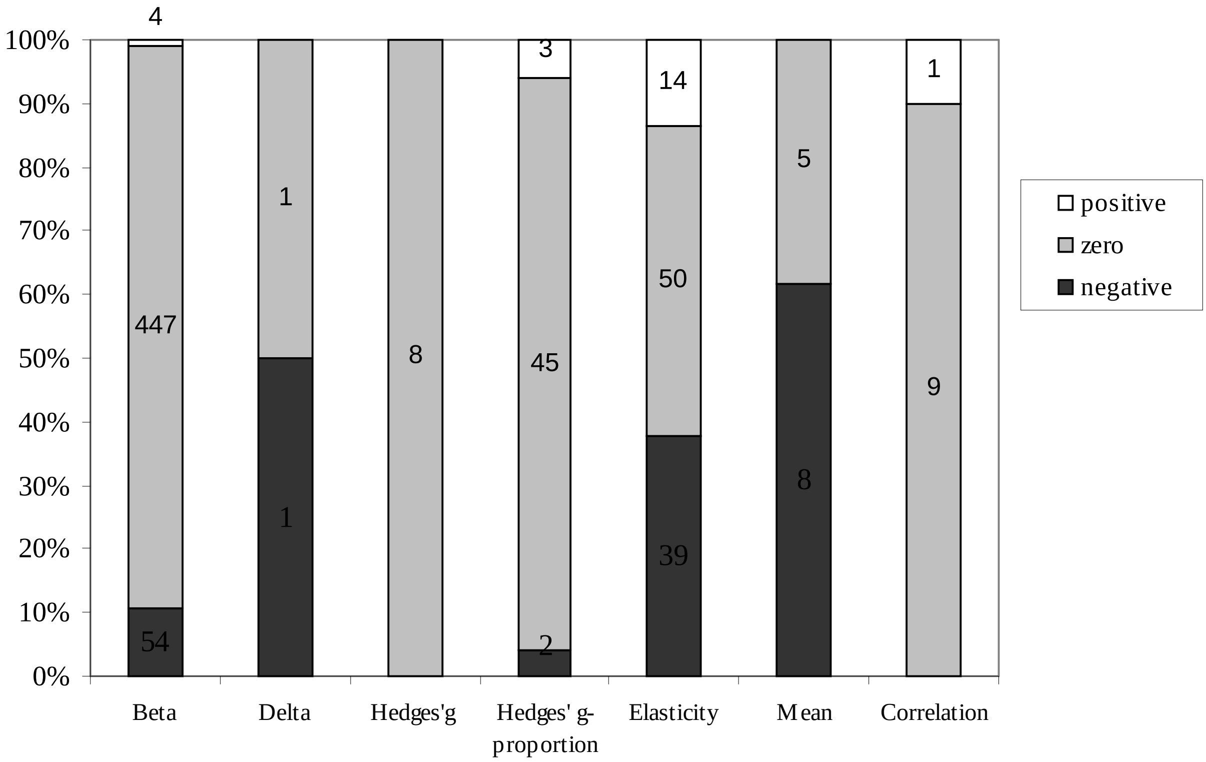

FIGURE 1. B. Percentage and number of negative, zero, and positive effect sizes according to effect size measure.

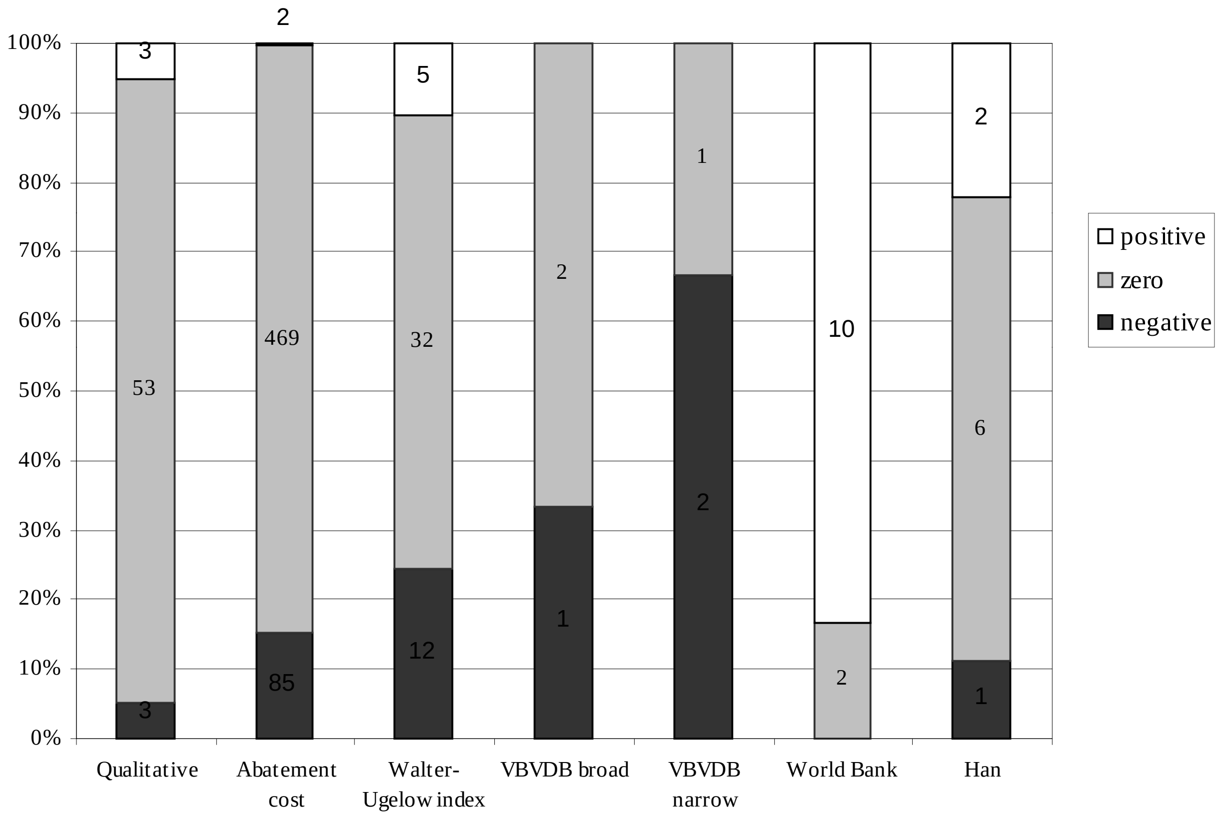

FIGURE 1. C. Percentage and number of negative, zero, and positive effect sizes according to type of stringency.